

UCB Twitter Course Final Project

Deliverables and Timeline

Project Subject Matter

The class project is your chance to decide what **you** want to work on and let your imagination run free and skills learned in this class and elsewhere shine! It is anticipated that projects will be one of the following, but you are free to suggest alternatives to these, subject to approval by the instructors:

- Software engineering for big data
- Big data analysis, including Twitter data
- Development of a software application using relevant data

Project Expectations

Students will work in teams of three. Students are expected to produce working software that explores one or more of the concepts that have been covered in class, either expanding on software engineering concepts, or analyzing one or a combination of data sources to gain new insight, or developing an application that makes use of relevant data. You will receive feedback on your project proposal to make sure it is neither too ambitious nor too narrow in scope. You may be asked to revise your project proposal before it is accepted. Proposals will be marked as either approved or not approved but will not be graded.

Time on Amazon EC2 clusters can be provided for class projects if the code has been tested in advance on an individual machine and shown to work. We will work this out on an individual project basis. It's a good idea to look at what resources are available on EC2; here is a helpful link: <http://aws.amazon.com/publicdatasets/>

Project Milestones and Timeline

All deliverables must be placed on the project web pages by the designated time to be considered turned in on time, with the exception of the proposal which can be turned in using the homework upload link.

- **[Due: Oct 22: 5:00 pm]** (*Not graded.*) Project proposal. This must include:
 - a) Team members, and their roles on the project.
 - b) Project goals, including what defines success (what should earn you an A, B, or C on the project)
 - c) Project timeline with milestones(Turn in via the same turn-in form we have been using for homework.)

UCB Twitter Course Final Project

Deliverables and Timeline

- **[Due: Nov 13, 9:00am]** **Note updated date.** (30% of the project grade) To keep you on track, there will be a mid-semester deliverable where you will be expected to be 30-40% done with your project. This deliverable will consist of:
 - a) A write-up consisting of the same material from Oct 22, a literature review of approaches to the problem that you are addressing; it must contain references to at least 9 publications or other references (3 per project member), as well as a summary of project goals and where you currently are,
 - b) Working code for a subset of your project, and
 - c) A class presentation describing project goals and where you are right now. Assuming 14 projects, that gives us 10 minutes/project. These will take place on **Nov 13 and 15.**
 - d) A description of what percent of work was done by each team member. (Turn in via your project web pages which will be set up by this time.)
- **[Due: Dec 10, 5pm]** (60% of the project grade) Final project due. This must include:
 - a) Project write-up consisting of the same material from Nov 12, updated as needed, and augmented with a full description of the results.
 - b) Documentation of any software developed. If an app was made, a link to the working app along with instructions for how to run it.
 - c) Description of which goals were met, or if goals changed, how and why this was done.
 - d) Software along with instructions to run it
 - e) Links to data and how to make use of it.
 - f) A description of what percent of work was done by each team member. (Turn in via your project web pages.)
- **[Dec 6 and Dec 11]:** (10% of the project grade) Final project presentations at Twitter and UCB, respectively.

Project Deliverables

There are several deliverables associated with the final project:

1. You are expected to maintain web pages describing your project, and its progress, on the class blog. If you want to keep it private to the class, that is ok, just let us know. You will be “turning in” the milestones on these web pages. You will also post links to results of analysis, web apps, software to run, etc, from these web pages.
2. You are expected to do a literature review of approaches to the problem that you are tackling. This can include papers from the scholarly research literature, well-informed blog posts, or other sources if relevant. The

UCB Twitter Course Final Project

Deliverables and Timeline

choices should be tailored to the work you are doing, be it data analysis or software development or software toolkits. You should write at least a paragraph about each document you review.

3. You are expected to write software to support your analysis or tool building. If you use code from other sources you must document this clearly.
4. You are expected to write high-quality prose describing your project goals, progress, execution, and any data analysis. Points will be taken off for poorly written write ups or write ups that are too terse.
5. You are expected to present your project publicly in a clear, brief manner. Presentations at Twitter will be 5 minutes long; those at UCB will be 10 minutes long.